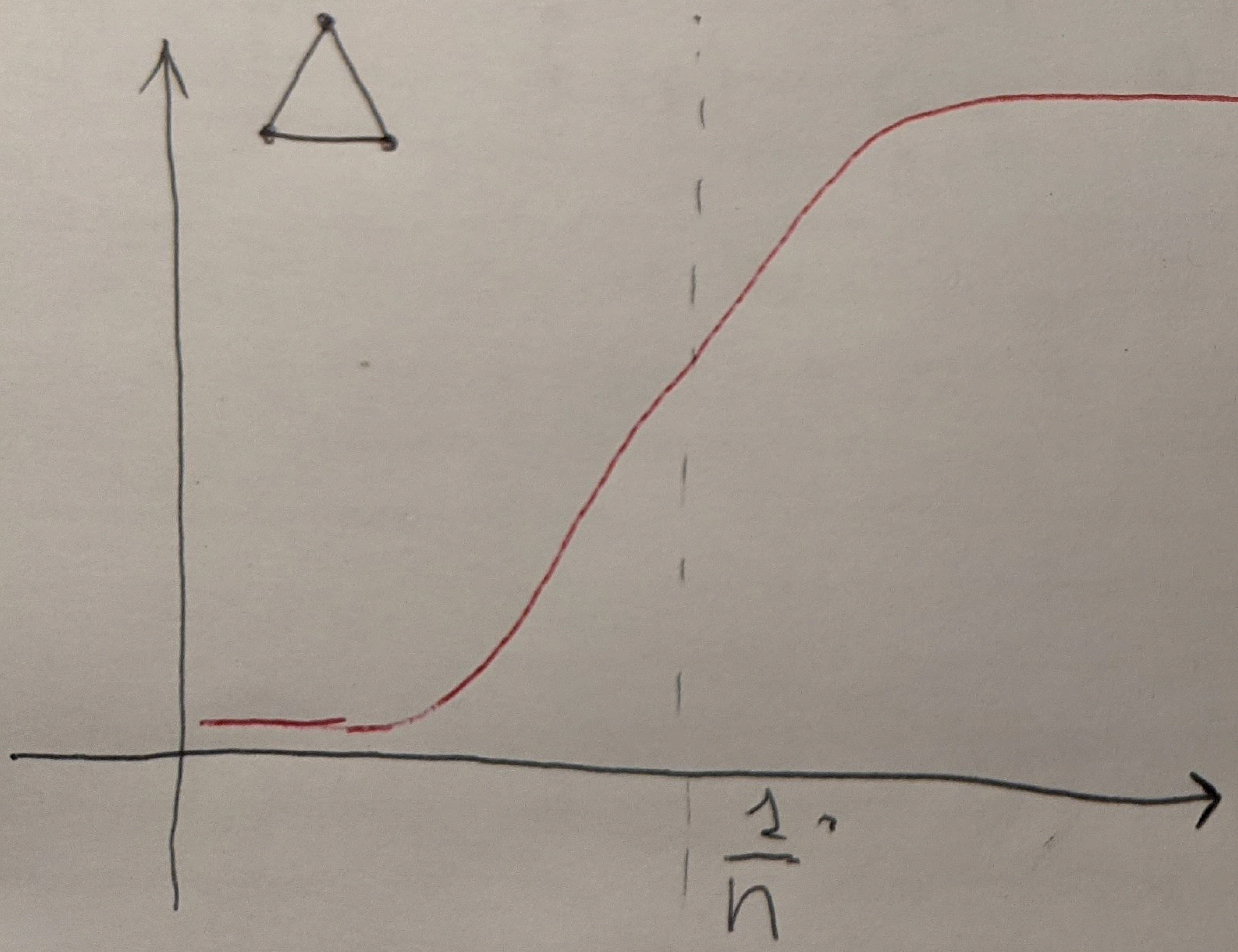
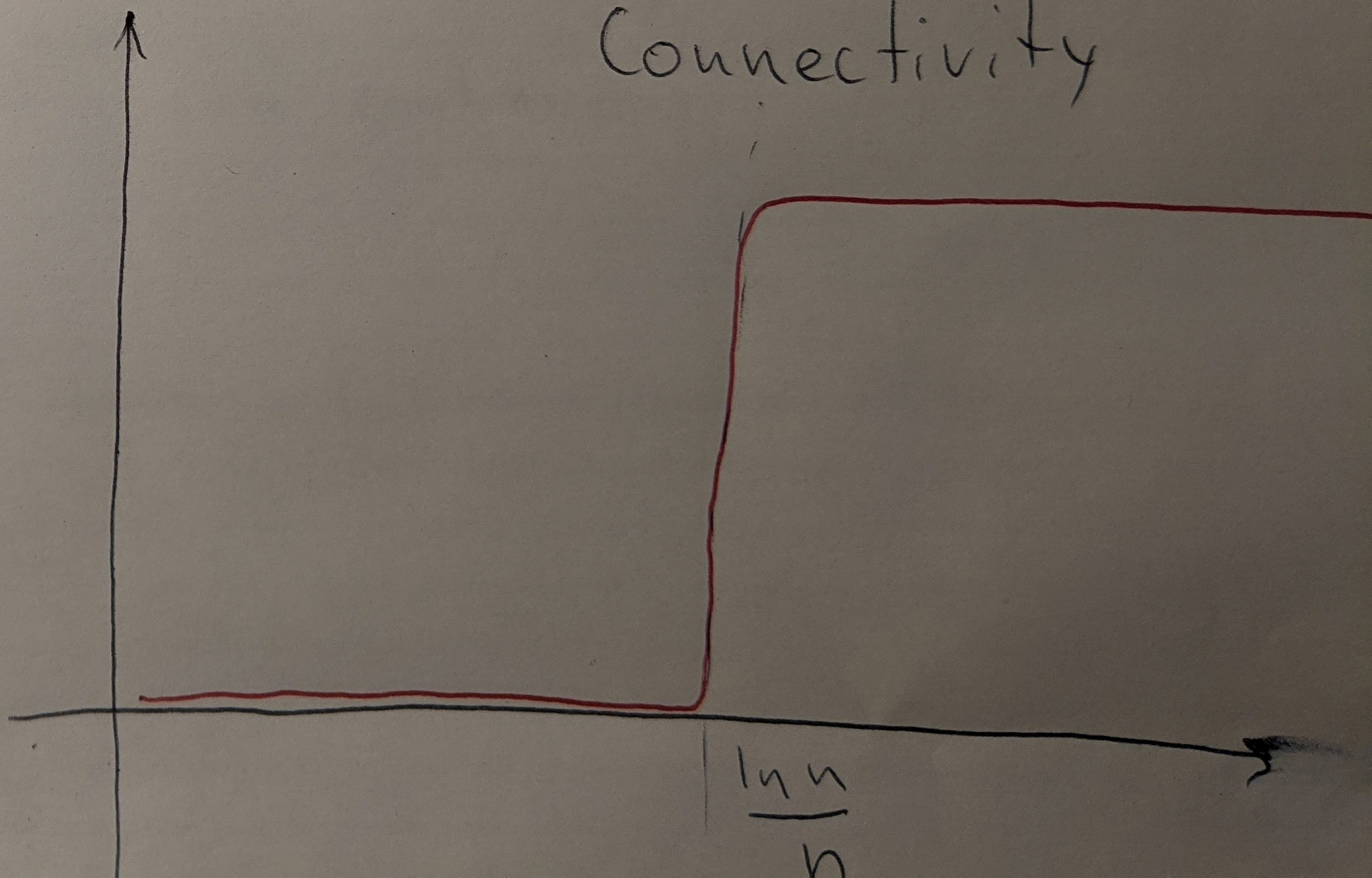
Lecture 8:

Existence of Thresholds & Chernoff Bounds





Previously:

$[n]_p$ - distribution on subsets of $[n] = \{1, 2, \dots, n\}$ obtained by selecting each element at random with probability p.

$\mathcal{P}([n])$ - collection of all subsets of $[n]$.

$\mathcal{F} \subseteq \mathcal{P}([n])$ is monotone if for $A \subseteq B \subseteq [n]$ if $A \in \mathcal{F}$ then $B \in \mathcal{F}$.

non-trivial if $\emptyset \notin \mathcal{F}$, $[n] \in \mathcal{F}$.

Lemma 4.6: $\mathcal{F} \subseteq \mathcal{P}([n])$ monotone & non-trivial (for fixed $[n]$) then $f(p) = \mathbb{P}([n]_p \in \mathcal{F})$ is continuous & increasing.

Proof: $f(p) = \sum_{A \in \mathcal{F}} p^{|A|}(1-p)^{n-|A|}$ → continuous.

Increasing → by coupling.

$r : [n] \to [0,1]$
$(r(1), r(2), \dots, r(n))$ each $r(i)$ is selected indep. uniformly in $[0,1]$

$[n]_p = \{i : r(i) \leq p\}$.

$p_1 < p_2$ selecting $A = \{i : r(i) \leq p_1\}$ $B = \{i : r(i) \leq p_2\}$.

generate (A, B) s.t. A is distributed according to $[n]_{p_1}$
B —″— $[n]_{p_2}$

and $A \subseteq B$.

$$f(p_1) = \mathbb{P}(\underset{=}{A} \in \mathcal{F}) \le \mathbb{P}(\underset{=}{B} \in \mathcal{F}) = f(p_2)$$

So $f(p)$ is non-decreasing.

With probability > 0 we have $A = \phi$, $B = [n]$,

$\to (p_2 - p_1)^n$

So $f(p_1) \le f(p_2) - (p_2 - p_1)^n \to$ so $f(p)$ strictly increasing.

Lemma 4.7: Let $\mathcal{F}$, $f(p)$ be as in 4.6.

then $1 - f(2p) \le (1 - f(p))^2$

Proof: Let A, B be random subsets of $[n]$ independently selected according to $[n]_p$.

How is $A \cup B$ distributed? $\boxed{[n]_{2p-p^2}}$

$$\mathbb{P}[i \in A \cup B] = 1 - \mathbb{P}[i \notin A]\,\mathbb{P}[i \notin B] = 1 - (1-p)^2 = 2p - p^2 < 2p.$$

$$f(2p - p^2) = \mathbb{P}(A \cup B \in \mathcal{F}) \ge \mathbb{P}(A \in \mathcal{F} \text{ or } B \in \mathcal{F}) = 1 - (1 - f(p))^2$$

$\wedge\!|$

$f(2p)$

$$f(2p) \ge 1 - (1 - f(p))^2 \quad \checkmark$$

Corollary 4.8: Every non-trivial monotone graph property has a threshold.
(Bollobás & Thomason) 1987

Proof: Need to show that there exists for each n
$p_c(n)$ s.t.

$$\mathbb{P}(G(n,p) \in \mathcal{F}) \to 1 \text{ as } p/p_c \to \infty \quad (*)$$
$$\mathbb{P}(G(n,p) \notin \mathcal{F}) \to 1 \text{ as } p/p_c \to 0. \quad (**)$$

Fix n. $f(p) = \mathbb{P}(G(n,p) \in \mathcal{F})$.

There exists p_c s.t. $f(p_c) = \frac{1}{2}$. by 4.6.

By 4.7.
$$1 - f(2^k p_c) \leq (1 - f(p_c))^k = \frac{1}{2^k}$$

$$f(\underbrace{2^k p_c}) \geq 1 - \frac{1}{2^k} \to 1 \text{ as } k \to \infty.$$

implies $(*)$

is in the other direction
$$1 - f(2^{-k} p_c) \geq (1 - f(p_c))^{\underbrace{1/k}} \geq \frac{1}{2^{1/k}}$$

$$f(2^{-k} p_c) \leq 1 - \frac{1}{2^{1/k}} \to 0 \text{ as } k \to \infty$$

implies $(**)$.

$\mathcal{F}$ = G has no isolated vertices.

$$(1-p)^n \sim e^{-pn}$$

\# isolated vertices ne^{-pn} $e^{\log n - pn}$

Threshold $\dfrac{\log n}{n}$

If $p = \dfrac{\log n + c(n)}{n}$ and $c(n) \to +\infty$

then $G(n,p) \in \mathcal{F}$ w.h.p.

and $c(n) \to -\infty$

$G(n,p) \notin \mathcal{F}$ w.h.p.

(Same for connectivity)

Sharp threshold

$\mathcal{F}$ has no sharp threshold if there exists $\varepsilon > 0$

and $p_c(n)$ s.t.

$$\mathbb{P}(G(n,p) \in \mathcal{F}) \in [\varepsilon, 1-\varepsilon]$$

for $p \in [(1-\varepsilon)p_c(n), (1+\varepsilon)p_c(n)]$.

K_3 containment or H containment has no sharp threshold.

but connectivity, existence of Hamiltonian cycles, have sharp thresholds.

Theorem 4.9: If monotone property $\mathcal{F}$ has no sharp threshold
(Friedgut 1999) then it can be approximated by a subgraph containment property.

[Another interesting topic: Thresholds vs. Expectation thresholds.]

5. Chernoff bounds

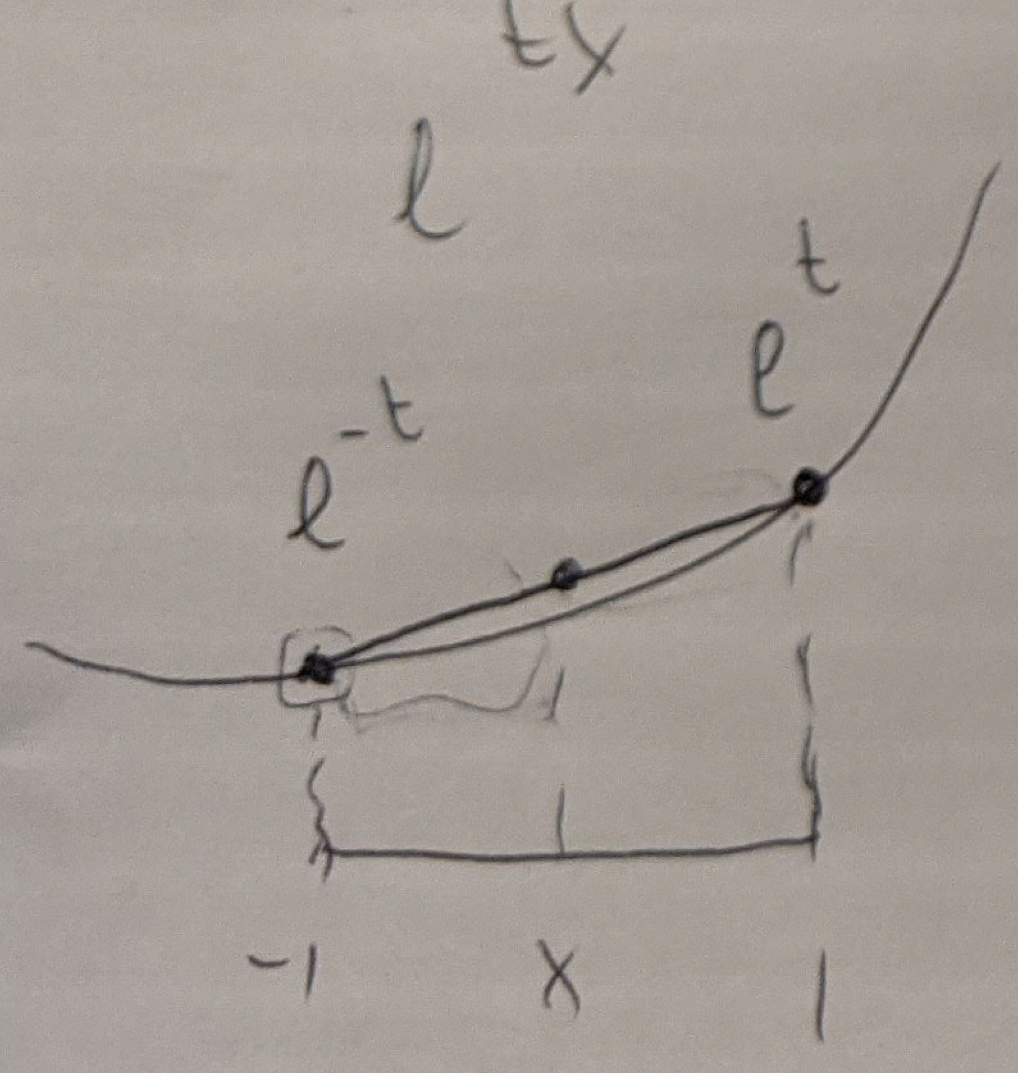
Theorem 5.1: Let $X = X_1 + X_2 + \dots + X_n$,
where $X_1, X_2, \dots, X_n$ are independent
$X_i \in [-1, 1]$, and $\mathbb{E}[X_i] = 0$.

$$\mathbb{P}(X \geq \lambda\sqrt{n}) \leq e^{-\lambda^2/2}$$

Proof: Moment generating function for X

$$t \to \mathbb{E}[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n]$$

$$\mathbb{E}[e^{tX}] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] \leq e^{\frac{t^2 n}{2}}.$$



$$e^{tx} \leq \frac{1-x}{2} e^{-t} + \frac{1+x}{2} e^{t} \quad \text{by convexity.}$$

$$\mathbb{E}[e^{tX_i}] \leq e^{-t} \underbrace{\mathbb{E}\left[\frac{1-X_i}{2}\right]}_{=\frac{1}{2}} + e^{t} \underbrace{\mathbb{E}\left[\frac{1+X_i}{2}\right]}_{=\frac{1}{2}}$$

$$= \underbrace{\frac{e^{t}+e^{-t}}{2} \leq e^{t^2/2}}_{\text{exercise?}}$$

$$\mathbb{P}(X \geq \lambda\sqrt{n}) = \mathbb{P}[e^{tX} \geq e^{t\lambda\sqrt{n}}] \underbrace{\leq}_{\text{Markov}} \frac{\mathbb{E}[e^{tX}]}{e^{t\lambda\sqrt{n}}}$$

$$= e^{\underbrace{-t\lambda\sqrt{n} + \frac{t^2 n}{2}}} = \underline{e^{-\lambda^2/2}}.$$

$$\frac{\partial}{\partial t} = -\lambda\sqrt{n} + tn$$

$$t = \frac{\lambda}{\sqrt{n}}$$

<u>Corollary 5.2</u>: Let $X = X_1 + X_2 + \dots + X_n$, $\mu = \mathbb{E}[X]$

X_i's are independent $X_i \in [0,1]$.

then $\mathbb{P}(X \geq \mu + \lambda\sqrt{n}) \leq e^{-\lambda^2/2}$.